

THE KAVERY ENGINEERING COLLEGE*(An Autonomous Institution)***M.E/MCA DEGREE END SEMESTER THEORY EXAMINATIONS, NOV/DEC 2025****Third Semester****Computer Science and Engineering****POE2404- Big Data Analytics****(Common to Master of Computer Applications)****Regulations - 2024****Duration : 3 Hrs****Maximum : 100 Marks****PART – A (ANSWER ALL QUESTIONS) (Marks 10*2=20)**

Q.No	Questions	BTL	CO	Marks
1.	Provide examples of industries or applications where big data analytics is used effectively.	L1	1	2
2.	Define Resampling.	L1	1	2
3.	Why it is important to have a good method of visualization?	L1	2	2
4.	What is ensemble analysis related to a large volume of data?	L1	2	2
5.	Classify Filtering stream with examples.	L2	3	2
6.	Generalize the surprise number (second moment) for the stream 3, 1, 4, 1, 3, 4, 2, 1, 2. What is the third moment of this stream?	L3	3	2
7.	Recall the components of Hadoop framework.	L1	4	2
8.	Recite the need for Apache pig.	L1	4	2
9.	How to deleting elements of a matrices and arrays.	L1	5	2
10.	Mention the various R Logical Operators	L1	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*16 = 80)

Q.No	Questions	BLT	CO	Marks
11.	a i Describe how the analytical scalability is handled in big data.	L2	1	8
	ii Examine the modern tools for big data analytics.	L2	1	8
	Or			
	b i Summarize in detail about the challenges of the big data in modern data analytics.	L2	1	10
	ii Explain the difference between the 'Big' part and the 'Data' part in Big Data, and discuss which one is considered more important.	L4	1	6
12.	a i Apply a genetic algorithm to solve a specific real-world problem. Provide an example and explain the approach.	L3	2	6
	ii Apply simulated annealing to solve a Traveling Salesman Problem (TSP) instance.	L3	2	10
	Or			
	b i Interpret the importance in classification of visual data analysis techniques.	L2	2	10
	ii Explain at least three fundamental visualization techniques and their use cases.	L2	2	6

13.	a	i	Analyze the data stream model by designing its architecture and explaining the role of each component using a clear and neat diagram.	L4	3	8
		ii	Examine the different types of motivation behind real time analytics platform and compare the various data analysis techniques available for time series data.	L4	3	8
	Or					
	b	i	Illustrate in detail about how to count the distinct elements in a stream.	L2	3	8
		ii	Generalize how data analysis is used in stock market predictions.	L2	3	8
14.	a	i	Demonstrate about HBase and Hbase clients in detail.	L4	4	8
		ii	Designate how Cassandra is integrated with Hadoop and also the tools related to Hadoop.	L4	4	8
	Or					
	b	i	List the advantages and disadvantages of NoSQL.	L4	4	6
		ii	Recommend a procedure to find the number of occurrences of a word in a document using Hive.	L4	4	10
15.	a		Consider a data frame with a = c(1, 2, 3), b = c(4, 5, 6), c(7, 8, 9) and find the value of the following	L3	5	16
		(i)	How do I select the c (4, 5, 6)?			
		(ii)	How do I select the 1?			
		(iii)	How do I select the 5?			
	Or					
	b	i	Demonstrate about Input-Output functions with a program.	L3	5	8
		ii	Write an R program to create a vector and to access elements in a vector.	L3	5	8